Technology Transfer Activities within the Federal High Performance Computing and Communications Program

April 1993

National Coordination Office for HPCC

Executive Office of the President
Office of Science and Technology Policy

Background and Overview

The High Performance Computing and Communications (HPCC) Program is a Federal initiative to promote the development and deployment of advanced computing and communication technologies. Goals of the program include:

- Extend U.S. technological leadership in high performance computing and computer communications
- Provide wide dissemination and application of the technologies both to speed the pace of innovation and to serve the national economy, national security, education, and the global environment
- Spur gains in U.S. productivity and industrial competitiveness by making high performance computing and networking technologies an integral part of the design and production processes

Strategies employed to achieve the goals of the HPCC Program include technology infusion into the Program through collaborations with academic and commercial sector partners; development of new technologies within the Program; the evaluation, deployment and transition to new technologies; and transfer of technologies from the

Program to the private sector. The means employed to achieve these goals and strategies include:

- Cooperative research and development agreements (CRADAs) and formation of consortia with academic and industrial partners;
- Use of agency and HPCC Program assets including:
    - the Interagency Interim National Research and Education Network (IINREN), which is composed of participating Federal agency subnetworks, and
    - NSF Supercomputer centers and other agencies' laboratories and computing research centers to support coordinated, collaborative research and development programs, and exchange of information and technologies;
- Formation and funding for inter-agency testbeds;
- Procurement and evaluation of prototype and early production systems;
- Support of basic research, education, and the development of human resources; and
- Sponsorship and conduct of technical meetings and workshops, and publication of scientific and technical papers and reports.

The transfer of technology from research to development and to application can be a very slow process due to a number of barriers: high initial cost, inadequate software and systems, software and systems unfriendly to users, and lack of standards. The HPCC Program relies upon substantial industry participation and cooperative proposals to stimulate the growth of shared knowledge and capabilities; improve the rate of technology transfer; and identify those new technologies of greatest potential for the future. The strategy and means employed to achieve the goals of the Program vary among specific activities supported by the Program and specific technologies being developed.

The diversity of the means employed by agencies to support their clientele and the complementary interactions of the programs of participating agencies all contribute to the synergism and richness of the HPCC Program as a whole. Below are examples of activities supported by the Program that illustrate the means and strategies employed to involve academic and industrial partners and achieve technology development, evaluation, and transfer.

High Performance Computing Systems

Some agencies provide up-front research and development funds leading to the production of high performance computing systems, early markets, and product evaluation.

- Advanced Prototypes of Scalable Systems

  The Advanced Research Projects Agency (ARPA) has recently contributed development funds for the Intel Paragon and Thinking Machines CM-5. The National Security Agency (NSA) is currently providing funds to Supercomputer

Systems Inc. NSA has also recently provided funds to E-Systems for development of a petabit storage system (which is now being marketed).

- R&D for Future Systems

  National Science Foundation (NSF)-funded research has developed special purpose computers for graphics calculations. Computer assisted design (CAD) tools for performing logic synthesis currently used in industrial endeavors are based on concepts developed and refined in NSF funded research projects. The software industry has been the recipient of many NSF funded basic research products. Innovations in the areas of performance evaluation software tools and compiler design have been instrumental in accelerating the development and use of scalable parallel systems. Efficient logic-based programming environments and simulation testbeds provide new, cost-effective means for software development.

  The National Institute of Standards and Technology (NIST) designed a high-performance, non-perturbing performance measurement instrument that is realized as a VLSI circuit. This instrument enables performance measurements, data collection, and transmission of information collected in massively parallel computing systems to an external computer system via an internal network interface. After NIST offered the design to U.S. industry, Intel Corporation incorporated it into its Paragon machine where the VLSI chip is used to collect system performance and utilization data which is used by the operating system to manage system resources.

  The National Security Agency (NSA) is an affiliate of MCC (the Microelectronics and Computer Technology Corp.) and within this consortium is a partner in Project CARNOT, which is building a system for transparent, seamless access to heterogeneous databases.

- Early systems evaluation

  Traditionally, NSA, the National Aeronautics and Space Administration (NASA) and the Department of Energy (DOE) have provided early markets for the latest supercomputer systems (e.g., Crays).

  Technical staff have provided feedback for supercomputer manufacturers from early product evaluations and have suggested product improvements for supercomputers that have resulted in product enhancements (e.g., hardware and instruction set modifications for Crays, programmable chip enhancements to the CM-2). NSA has recently made significant technical comments on the instruction set and performance choke points of Supercomputer Systems Inc. machine.

Advanced Software Technology and Algorithms

High performance algorithms have been the focus of many research efforts. A broad spectrum of new algorithms have been formulated that are currently being applied to such diverse industrial and commercial interests as linear programming and optimization problems; airline reservation systems and other on-line system development; pattern design in the textile industry; and cryptographic analysis.

- Portable, Scalable Software Libraries

  NIST developed a taxonomy for the classification of mathematical software that has been widely adopted. This taxonomy provides a framework for a software-based system called "Guide to Available Mathematical Software" (GAMS), which was developed by NIST for the classification and exchange of high quality, public domain software. GAMS allows a user, located somewhere on a network, to make a query about software satisfying his or her needs, select software satisfying those needs, and then to have the software and/or its documentation retrieved from a repository and delivered to the user's computer over the network (e.g., the National Research and Education Network (NREN)). The GAMS system was contributed to the NASA coordinated Software Repository and Exchange Experiment of the HPCC Program that is described below. As such, GAMS is part of the infrastructure supporting technology transfer within the Program.

- Software Repository and Exchange Experiment

  The goal of the High Performance Computing and Communications Software Exchange (HPCC-SE) is to facilitate the exchange and reuse of software. Its specific objectives are to: develop and demonstrate a distributed architecture and supporting technology that support software exchange; implement an initial distributed HPCC-SE that supports the needs of the HPCC; and specify an open non-proprietary architecture that will facilitate the emergence of a national software exchange. The development will be done in conjunction with a community-based Open Architecture Working Group.

  The development of the initial software exchange capability is scheduled to take place in three two-year phases. An experimental system will help discover some of the intrinsic problems of distributed software sharing and is intended for a limited audience. A prototype system will be a more extensive version and will serve a much wider audience. The final stage will be an initial operational system, which will provide the initial operational capability of the HPCC-SE. The results of the experimental system will generate requirements for the Prototype system. Similarly, the prototype system will provide requirements for the operational system.

The Software Exchange Experimental System is currently accessible over the Internet. Internet is viewed as being part of one large "logical library" in which all databases (e.g., software repositories), and search mechanisms (e.g., repository directories, catalogues, etc.) appear as items on the "shelves" of this library.

- Software Tool Development

  The Environmental Protection Agency's (EPA) technology transfer program targets the dissemination of high performance computing and advanced environmental modeling capabilities to select groups from industry, states, and Federal users. The industrial group is represented by the Consortium for Advanced Modeling of Regional Air Quality. The target groups of state users are selected from states with professionals involved in front-line environmental decision making and using models of regional and urban air quality.

  Within the EPA HPCC program, technology transfer is approached as a two-stage process. The first stage is technology infusion to HPCC R&D teams. This infusion includes coordination and cooperation with the development-oriented agencies within HPCC, such as DOE and NASA, to move the technology rapidly to the more mission-oriented agencies, such as EPA. The second stage is dissemination of high performance computing modeling capability to front-line environmental professionals. The second stage is conducted as an identified, organized technology transfer program component within EPA's HPCC program. Users will gain first-hand experience about how high performance computing can alleviate many of their decision making constraints while at the same time they will work with developers to ensure that system capabilities and interfaces address their needs.

  Training and feedback on functionality and interfaces will occur in parallel with prototype development cycles. A high performance computing learning environment is created through workshops, involvement of professional trainers and curriculum developers, and establishment of training facilities at a regional supercomputing center. The program will test approaches to the transfer of high performance computing capabilities to less-sophisticated front-line environmental modelers and will develop infrastructure requirements to support a distributed base of users of environmental models for decision making across the country. The IINREN is an important HPCC Program asset influencing EPA's program. Collaborative programs with universities and other supercomputer centers are also used to help achieve the technology transfer goals.

  The NSF Supercomputer centers Program provides scientists and engineers with access to advanced computing resources allowing early use of emerging architectures in scientific research. The centers also train researchers and students in the use of those resources; develop new software tools and systems

to increase the productivity of the computational research community; identify the research needs of the computational research community to computer and software vendors; and provide numerous mechanisms for collaboration between academic and industrial researchers. The primary support vehicle to accomplish program goals are long-term Cooperative Agreements, which support the operational and capital costs for four National Supercomputer centers. Within the Cooperative Agreements it is specifically stated that the centers are encouraged to submit proposals to industrial organizations as appropriate, and all the centers have availed themselves of this opportunity.

The centers do interact with their industrial partners as collaborators and not merely as providers of "computing cycles." Many private sector organizations have little experience with high performance computing, thus, affiliation with a national center offers an innovative method of exploring and ultimately exploiting the usefulness of this technology. There are however, some very experienced users of computational methods in industry who have come to the centers for the intellectual environment they offer. Each partner brings specific problem needs that can be solved by center staff or affiliated faculty. The New York-based banking company J.P. Morgan has recently joined the industrial supercomputing program at the National center for Supercomputing Applications (NCSA). There is now a group of eight major corporations working at the center: AT&T, Caterpillar Inc., FMC Corp., Dow Chemical Co., Eastman Kodak Co., Eli Lilly and Co., J.P. Morgan, Motorola Inc. and Phillips Petroleum Co.

Examples of industrial involvement at the NSF Supercomputer centers include the following:

- Researchers at Corning Glass have been working with scientists at the Cornell Theory center for many years developing methods of simulating atomic interactions to better understand and utilize the properties of glass. They have achieved increases in computational speeds approaching 1,000, which enables more accurate studies of insulators and semiconductors.
- Xerox resident scientists collaborate in the research groups of faculty in several departments at Cornell o mechanical and aerospace engineering, computer science, and civil and environmental engineering. This pre-competitive research studies means of shortening the time from conceptual initiation of a project, through the design phase to final product appearance. Other companies are being solicited to join the program, which is housed in the Theory center and which makes use of Theory center resources.
- Phillips Petroleum is a new industrial partner at the National center for Supercomputing Applications (NCSA). The company is a very sophisticated user of computational techniques, and probably has as much or more computer power available internally as do all the NSF centers combined. Yet, Phillips believed that an affiliation with NCSA would allow them to gain experience with visualization, and have asked NCSA's staff to examine the networking of all the worldwide Phillips computers.
- The first pharmaceutical company in the country to purchase a supercomputer was Eli Lilly, following a three-year partnership with NCSA. That collaboration allowed Lilly to gain knowledge on how to design and develop new drugs by

computational means. However, even after acquiring their own Cray 2, Lilly returned to NCSA as a partner in order to investigate new computer systems coming online at the center.

- With 80 billion soft drink cans used yearly, any means of lessening the cost of production can be very valuable. By cooperating with the Pittsburgh Supercomputing center, Alcoa has changed its design methods to include computational techniques. Methods have been developed to design a can to use the thinnest aluminum that will still meet specifications for strength and appearance; and design time has been reduced from months to days.
- U.S. Steel has turned to the Pittsburgh Supercomputing center in order to understand many aspects of the steel production process. By using a supercomputer to compute the solution of a finite element model of hot steel rolling, high quality steel products can be produced at a competitive cost. They are also investigating ways to improve the process control systems used in steel manufacturing.
- Scientists at GenCorp Research wanted to design a rapid production process for compression-molding of reinforced plastic auto body panels. They used the San Diego Computer center's Cray to design optimum molding patterns for the panels of several auto bodies, including the Chevrolet minivans. The process is working at GenCorp's giant press in Shelbyville, IN. The supercomputer has permitted high turnaround for an iterative design process, essential in planning and production.

The center for Research on Parallel Computation (CRPC) is a NSF Science and Technology center. Participating institutions are Rice University, California Institute of Technology, Los Alamos National Laboratory, Argonne National Laboratory, University of Tennessee, and Syracuse University. Below are selected technology transfer activities stimulated by the CRPC.

- Thinking Machines has incorporated communications routines in their CMSSL product that were derived directly from the CRPC differential equations work.
- The NEKTON code developed by Paul Fisher of the Caltech CRPC group and a group at the Massachusetts Institute of Technology under A. Patera is being turned into a commercial package for solving very general Navier-Stokes problems in incompressible flows. A company called NEKTON has been formed as a result.
- Caltech hosted a workshop for 120 industry representatives to expose them to high-performance computers and to encourage them to take advantage of this resource. Meetings have taken place with TRW, Ford Motor Co., McDonnell Douglas, Lockheed, and other industry representatives (separately) to discuss the use of parallel computers at their sites.
- In November 1990, CONVEX Computer Corp. announced its new applications compiler, giving substantial credit to the compiler group at Rice for ideas underlying the system. The Convex applications compiler analyzes and transforms whole programs to eliminate errors and improve execution efficiency. The underlying implementation technology, called "interprocedural analysis and optimization," was first implemented in a practical compilation system in the ParaScope programming environment by researchers at Rice.
- The IBM XL series of compilers for the RS/6000 uses register allocation techniques developed at Rice.
- CRPC researchers at Caltech have been working with scientists at Intel Supercomputer Systems Division since 1990 on improving the hardware and software environment for the Delta and future Intel products. As a result, many changes have been made to Paragon software and hardware -- key areas affected are high-speed input/output interfaces, concurrent file systems, graphics library

contents, file back-up hardware and software, and global communications functions.
    - CRPC researchers at Caltech and Los Alamos National Laboratory have been working closely with representative from Kodak Datatape Division to design a high-performance parallel interface (HIPPI) for their ID-1 tape drives, and the software to control it.
    - The High Performance FORTRAN Forum (HPFF), coordinated by the CRPC, is a coalition of industrial and academic groups working to identify the basic issues and suggest a set of standard extensions to FORTRAN to provide the necessary information. The HPFF includes most vendors currently delivering parallel machines; government labs; and many university research groups. Since its introduction over three decades ago, FORTRAN has been the language of choice for scientific programming for sequential computers. Parallel computing, however, has no such standard, leading to a software crisis on these high-performance machines. This crisis results from the lack of features applicable to several areas, including opportunities for parallel execution; type of available parallelism MIMD/SIMD (multiple-instruction multiple- data versus single-instruction single-data), or some combination of the two; allocation of data among individual processor memories; and placement of data within a single process. The intent of the HPFF is to develop extensions to FORTRAN that provide support for high performance programming on a wide variety of machines, including massively parallel SIMD and MIMD systems and vector processors.

National Research and Education Network

- Gigabit R&D Activities

  ARPA, Naval Research Laboratories (NRL), NASA, and NSA have cooperated in developing secure, gigabit network testbeds. The joint NSF/ARPA gigabit testbed project administered by the Corporation for National Research Initiatives (CNRI) has many industrial players. These include AT&T Bell Labs and Bellcore, AT&T Long Lines, MCI, six of the seven Regional Bell Operating, IBM, and others. NSF and ARPA funding for the project is $15.8 million; estimates of the investment by the private participants are between $100 million and $300 million.

  Jointly, DOE and NASA are completing a "technology procurement" contract for early development and deployment of broadband circuit-switched technology. This procurement will complement the installed base of packet-switched technology in the IINREN and provides an opportunity to push technology development and evaluate an alternate technology. This technology is closely related to that envisioned for future deployment by the telecommunications industry for Integrated-Services Digital Network.

- Interagency Interim NREN upgrades/deployment of new technology

  An industry seeded by the NSFNET is that of IP packet switches/routers. NSF's standardization on the TCP/IP protocol suite in late 1985 solidified and organized the market for IP hardware. Although Cisco and Proteon predated the NSFNET program, Network Systems entered the market and Wellfleet was started in

direct response to the NSF-catalyzed demand. It is perhaps noteworthy that the demand for IP switches is worldwide, but the manufacturing base is entirely U.S.; IP networks in Europe, for example, (which have an estimated two million users) are implemented almost entirely with Cisco equipment.

The NSFNET program has been directly responsible for the emergence of at least three of the commercial providers of Internet connectivity and access: PSI from NYSERNET, ANS from the NSFNET Backbone, and CERFNET from SDSCNet. Collaborations exist in some, though by no means all, of the regional networks in the NSFNET system. Two examples are NYSERNET and PREPNet. NYSERNET in the state of New York was begun as a cooperative project with NYNEX (one of the "Baby Bells"), although that association has since changed. The Pennsylvania Research and Economic Partnership Network (PREPNet) is a collaboration between the University of Pittsburgh, Carnegie-Mellon University, the Commonwealth of Pennsylvania and the Bell Telephone Company of Pennsylvania.

Summary and Conclusions

A diverse set of activities and means are used to develop and disseminate technology associated with the HPCC Program. In addition to these specific activities, lectures, workshops, conferences and other technical forums are sponsored to involve researchers in academia and the private sector. These forums and their proceedings, scientific and technical papers, reports and journals provide other means to advertize and disseminate the results of the Program. The richness of the activities and diversity of the means used to engage the research and development communities by participating agencies is a strength of the Program. It provides opportunities for a broad spectrum of participants and the diversity minimizes the risk of failure of a single research and development option. The active engagement of R&D partners from industry and academia insures technology transfer through active participation (e.g., consortia, CRADAs) and wide dissemination through publications and activities designed for public sharing of products of the Program (e.g., software repositories).